Integrating Vision Capabilities into HOAP-2 Humanoid Robot

Jiajing Xu, California Institute of Technology

Mentor: Dr. Adrian Stoica, Jet Propulsion Laboratory

Co-mentor: Dr. Didier Keymeulen, Jet Propulsion Laboratory

September, 2004

Abstract

This paper describes the work done to provide humanoid robot HOAP-2 with fundamental vision capabilities. These features, when integrated with motion software, allow the robot to track and reach target objects, fetch them and transport them to the destinations in its visual field.

Particular image processing steps include grabbing frames from computer memory, detecting distinctive feature, calculating centroid, estimating distance, and guiding the robot walking towards the object. The decision made upon the above analysis is passed to the real-time control module on the robot. This paper also briefly overview the system and software architecture for operating HOAP-2.

Introduction

Limited to work too tedious or dangerous for humans, today's robots weld parts on assembly lines, inspect nuclear plants, and explore other planets.[1] Generally, robots are still far from achieving their fictional counterparts' intelligence and flexibility. Humanoid robotics labs worldwide are working on creating robots that are one step closer to science fiction's androids. The key reason for preferring humanoids is their optimal shape for being taught by humans and learning from humans.[2]

Autonomy often represents a trade-off between performance on particular tasks and generality in dealing with a broader range of stimuli.[3] Unlike industrial robots that operate in a fixed environment on a small range of stimuli, humanoid robots must operate flexibly under various environmental conditions and for a wide range of tasks. Building autonomous systems provides robustness and flexibility that task-specific systems can never achieve. Although the environment will not be nearly as hostile as those planetary explorers face, they are also not tailored to the humanoid robots. Vision is an ideal sensor modality for intelligent robots. It provides rich information on the environment as required for recognizing objects and perceiving environment in real time. Moreover, vision-guided humanoid robots may be largely calibration-free, which is a great practical advantage.

In this paper, I present the process of integrating vision capabilities to the second generation of the Fujitsu's Humanoid for Open Architecture Platform (HOAP-2). When integrated with my teammate's motion library, the robot perceives the environment and is able to locate and track an object. HOAP-2 is equipped with 25 servo actuators: 6 for each log, 4 for each arm, 1 for each hand, 2 for the head, and 1 for its waist. As for feedback aspect, there are 4 pressure sensors on the bottom of each foot, and an accelerometer and gyroscope inside the torso. The vision system consists of two CCD cameras, capable of capturing frames of 640 by 480 pixels at a fairly low speed. The images are sent to an Intel Pentium 4 2.4 GHz computer via USB communication. After image processing, the main program uses the outcome to generate a series of motion sequences, which will be sent back to the HOAP-2 robot. The on-board motor encoder of HOAP-2 implements the low level motor mechanics. (See in Figure 1 System Diagram.)

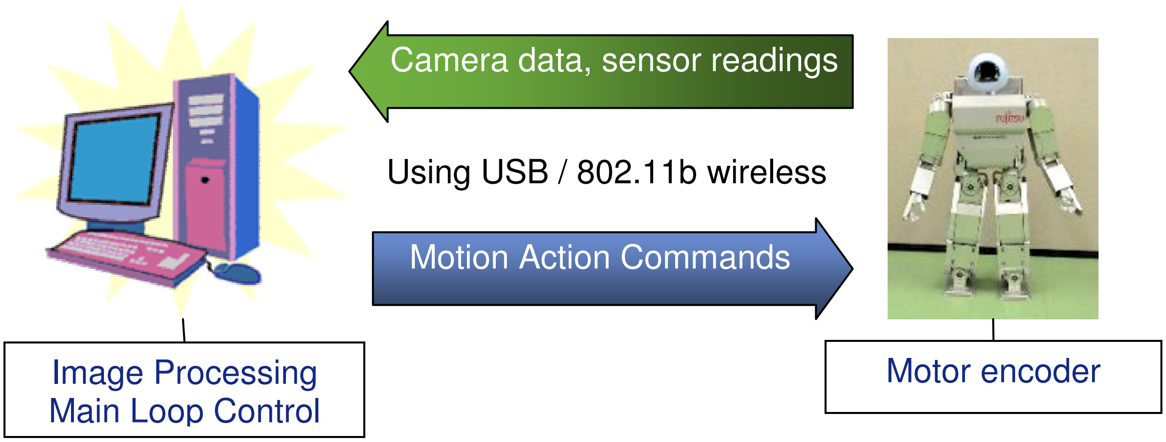


Figure 1 System Architecture

Programming on HOAP-2 is a bit different from usual programming. The real-time programs communicates with the user program (non real-time) via a special block of memory, called Shared Memory. (See Figure 2 Programming Architecture) User has to create a First-In-First-Out to access the shared memory.

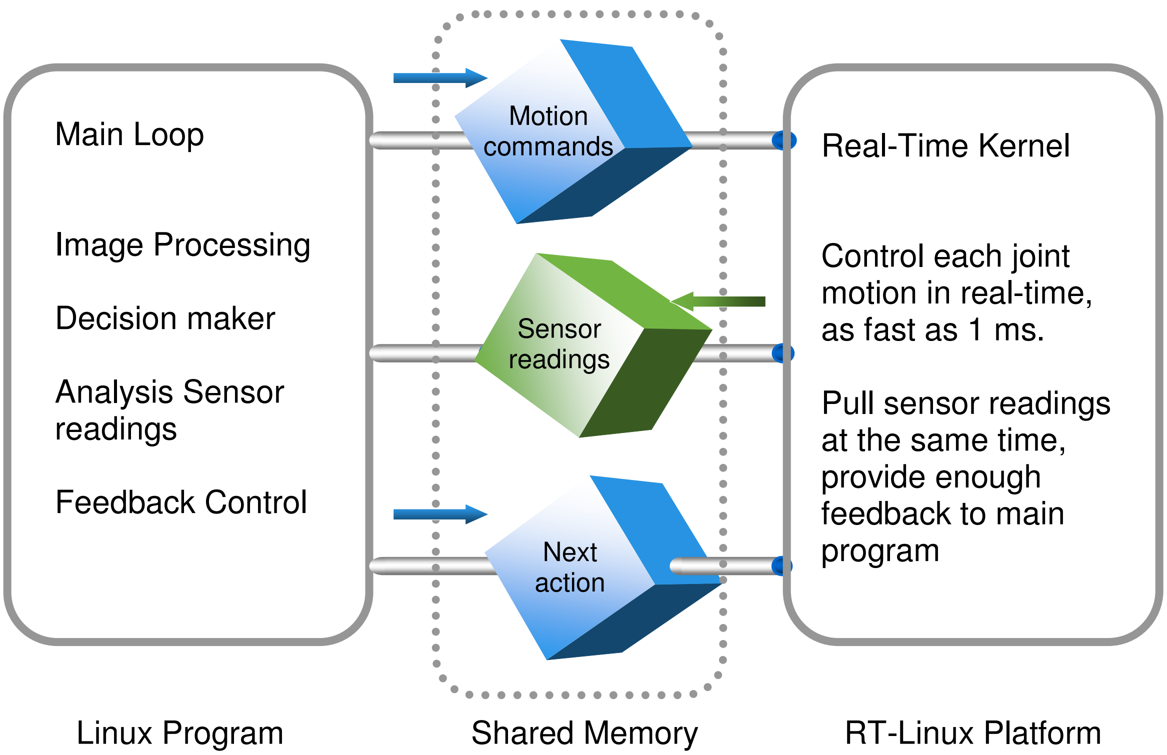


Figure 2 Programming Architecture

Findings

I chose red color for the target and green color for the destination instead of making a perfect white background and all black targets in the lab. There are two phases in the image processing part. Firstly, I apply color filter and noise-cancellation filters to the raw image frames captured from both cameras. Then, using the processed data, I can obtain the position of the 2D centroid on each frame. Since two cameras view the object from different angle, I derived a formula to map the difference to the distance from the object to the robot. With these data, I can also indicate if the robot is facing towards the object. The latter is extremely useful when the robot is trying to reach the object. The robot can maintain itself walking towards the object by turning a small angle on the run.

With programming goes on, I found out that the distance estimation part is a little tricky. The trigonometry formula I obtained requires two parameters from the camera, which is not documented by the manufacturer. I designed several experiments to measure them successfully to ensure the accuracy of the future calculation.

Discussion

Results

After intense testing in the lab, the vision-guide walking is proven to be perfectly functional. The robot, when integrated my teammate's motion library, can walk following the object provided that it is moving slowly than the sample rate of the vision system. On the other hand, the distance estimation is not as ideal as I proposed. The results are good within the range from 20 cm to 200 cm. If the actual distance is beyond that range, the error percentage from the image processing increases rapidly. I did extra experimental measurements to come up with compensation functions to bridge the gap between theoretical results and experimental readings. A slight improvement is observed.

The Future

For the future development of the vision system, it is essential to use a better image processing algorithm to analyze the environment, such as object recognition and distance estimation from two cameras. In this case, color segmentation is an alternative solution of using the color filter as I did. Furthermore, an additional processor, such as a separate FPGA or DSP, is helpful to increase the sample rate of the camera. The high speed processing ability ensures that more complex image processing and more accurate algorithms could be performed in real time. This provides the robot more accurate and important information. With these rich results, it is possible for the robot to analysis the terrain, and avoid obstacle without a stereo-vision system in real-time. The next promising goal for this research would be to enable HOAP-2 to digitalize the 2D schematics and reconstruct 3D structures from it. It will be interesting to see the interaction with human being when face recognition is integrated as well.

Methods

Process Raw Image from Camera

This section presents the simple image processing done in this project. The vision system receives both left and right images from the CCD cameras in the robot head. Both cameras are calibrated carefully and focal length and principal points are measured. These parameters are important for the following *Vision Computation* section.

Firstly, in order to filter colors other than red or green, the following RGB rules are used.

$$\sqrt{(R-255)^2+G^2+B^2} \le \frac{255}{2} \text{ , for red color}$$

$$\sqrt{R^2+(G-255)^2+B^2} \le \frac{255}{2} \text{ , for green color}$$

In addition, the following 3X3 neighborhood operators are applied to the raw images so as to increase the accuracy of the future calculation.

-1	-1	-1
-1		-1
-1	-1	-1

Noise Cancellation

1/6	4/6	1/6
4/6	-20/6	4/6
1/6	4/6	1/6

Edge Detection

In the testing phase, this method is proven to be working. Screenshots are as follows:

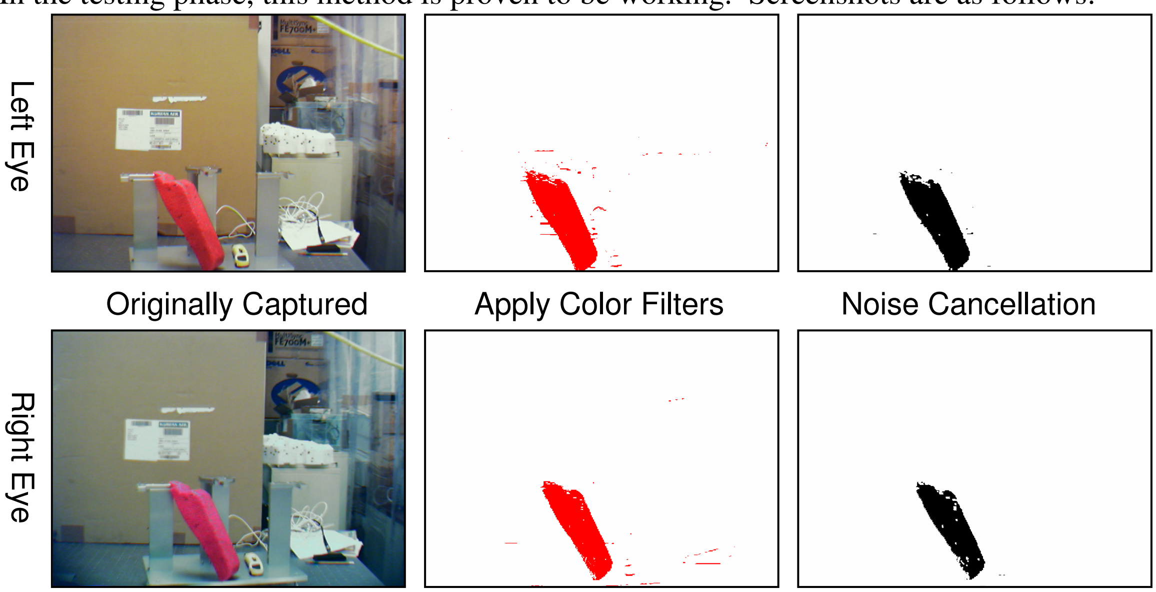


Figure 3 Image Processing Screenshots

Vision Computation
Given the processed image from previous section, I will perform a series of computation in this section. These features give the robot the capability to track the object and estimate the distance away from the object.

First of all, using the definition of the centroid in physics, I locate the 2D centroid of the object in each frame. Denote the horizontal offsets of the centroids are *d1* and *d2* for left and right camera respectively. In Figure 4, I relate the offsets of the centroid to the distance from the object to the robot.

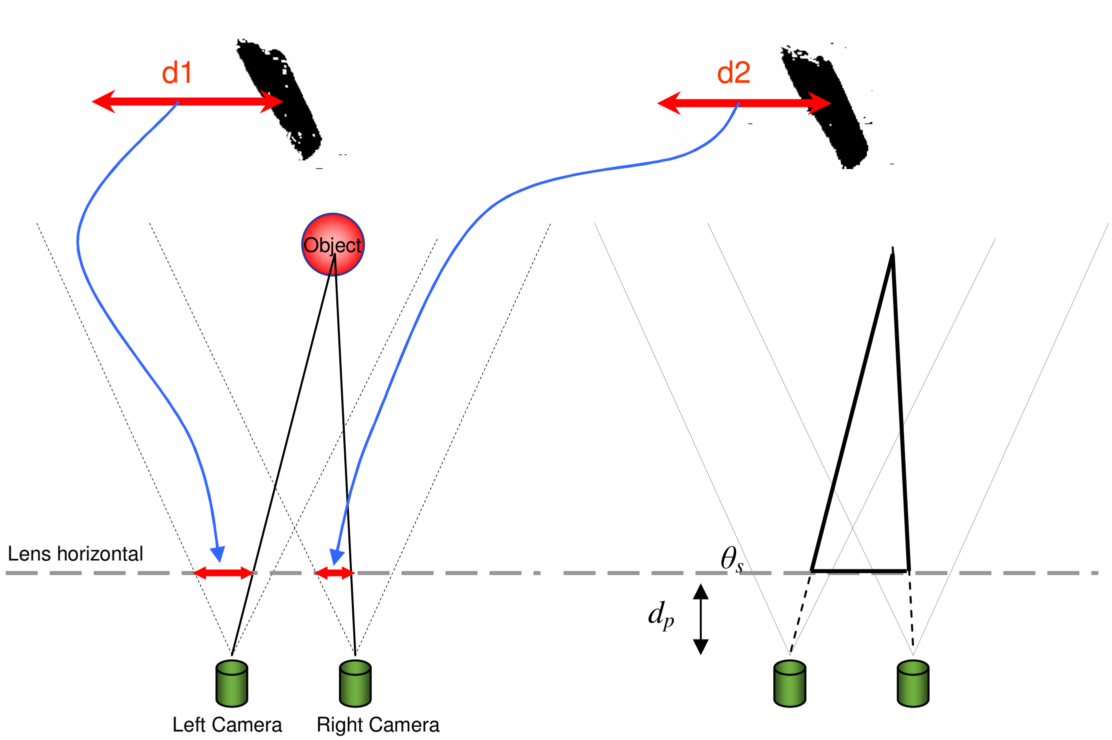


Figure 4 Vision Computation

The constant angles and lengths used are defined as follows:

$d_p = 9.45cm$

$d_g = 6.0cm$ *(distance between two cameras)*

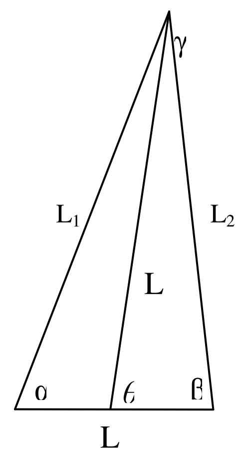
$\theta_0 = 50.5$ *deg (viewing angle of two cameras)*

$\theta_s = (180 - \theta_0)/2 = 64.75$ *deg (see figure 4)*

In the key triangle in the right part of Figure 4, two side angles can be obtained using trigonometry identities. They are,

$$\alpha = \frac{d_p}{\tan\theta_S} \Bigg/ \frac{d_1 - 320}{640},$$

$$\beta = \frac{d_p}{\tan\theta_S} \Bigg/ \frac{d_2 - 320}{640}$$



$$l_1 = l_3 \cdot \frac{\sin\beta}{\sin\gamma}$$

$$l_2 = l_3 \cdot \frac{\sin\alpha}{\sin\gamma}$$

$$l = \sqrt{l_1^2 + l_2^2 - \frac{l_3^2}{4}}$$

$$\theta = \cos^{-1}\left(\frac{l^2 + \frac{l_3^2}{4} - l_2^2}{2 \cdot l \cdot \frac{l_3}{2}} \right)$$

Once the two side angles are obtained, $\gamma = \pi - \alpha - \beta$ by the law of inner angles in triangle. *l1* and *l2* can be obtained by using the law of sine; the distance from the center of the bottom side to the tip *l* and the angle θ, which indicates if the robot is facing the object, are obtained by the law of cosine.

If $\theta > \pi/2$, object is on the right hand side of the robot;
If $\theta < \pi/2$, object is on the left hand side of the robot;
If α, β, or θ is not defined, (since the narrow viewing angle of the cameras), robot has to turn its body to re-scan the object.

Acknowledgements

I would like to thank my mentor Adrian Stoica and my co-mentor, Dr Didier Keymeulen for giving me this opportunity to work on this project. In addition, I would like to thank my teammates James Jeanne and Tim Hidalgo for their contribution to this project.

Reference

[1] R.A. Brooks et al., Alternative Essences of Intelligence, *Proc. 15th Nat l Conf. Artificial Intelligence (AAAI 98) and 10th Conf. Innovative Applications of Artificial Intelligence (IAAI 98),* AAAI Press, Menlo Park, Calif., 1998, pp. 961—968.
[2] Stoica, Adrian, Humanoids for Urban Operations, *Jet Propulsion Laboratory White Paper,* Pasadena, Calif., 2004
[3] R.A. Brooks et al., Humanoid Robots: A New Kind of Tool, C. Nehaniv, ed., *Springer Lecture Notes in Artificial Intelligence, Vol. 1562*, Springer-Verlag, Berlin, 1998.